

**NON-UNIFORM NON-ASYMPTOTICAL SHARP
ESTIMATE FOR THE RATE OF CONVERGENCE
FOR BERNSTEIN'S POLYNOMIAL APPROXIMATION,**

with bilateral constant evaluation

EUGENE OSTROVSKY, LEONID SIROTA

Bar-Ilan University, 59200, Ramat Gan, ISRAEL;

e-mail: eugostrovsky@list.ru

e-mail: sirota3@bezeqint.net

ABSTRACT.

We derive the non-asymptotical non-uniform sharp error estimation for Bernstein's approximation of continuous function based on the modern probabilistic apparatus.

We investigate also the convergence of derivative of these polynomials and we will consider briefly also the multivariate case.

Key words and phrases. Bernstein's polynomials, continuous function, approximation, norm, tail of distribution, Grand Lebesgue Spaces and norms, bilateral constants evaluation, Binomial (Bernoulli) distribution, subgaussian, strictly subgaussian and almost strictly subgaussian random variables, modulus of continuity, binary and Rademacher's r.v., derivatives, multivariate case, sharp estimation, Hölder and Lipschitz conditions, trial functions.

1 Introduction. Notations. Statement of problem.

Let $f = f(x)$, $x \in [0, 1]$ be continuous: $f \in C[0, 1]$, numerical function, $n = 2, 3, \dots$,

$$B_n[f](x) = \sum_{m=0}^n \binom{n}{m} f\left(\frac{m}{n}\right) x^m (1-x)^{n-m} \quad (1.0)$$

be its Bernstein's polynomial of degree n .

It is reasonable to define formally

$$B_\infty[f](x) = f(x), \quad f(\cdot) \in C[0, 1]. \quad (1.0a)$$

Denote also as ordinary

$$\omega[f](\delta) = \sup_{x, y \in [0, 1], |x-y| \leq \delta} |f(x) - f(y)| \quad (1.1)$$

module (modulus) of continuity of the function $f = f(x)$ of a first order,

$$\Delta_n[f](x) = |B_n[f](x) - f(x)| \quad (1.2)$$

be an *non-uniform* error of the Bernstein's approximation of a n^{th} order for the function f .

The problem of error estimation for Bernstein's approximation of continuous function goes back to the classical work belonging to S.N.Bernstein [3] (1912). He proved in the famous article [3] that as $n \rightarrow \infty$ the sequence $B_n[f](x)$ of a Bernstein's polynomial of degree n converges uniformly as $n \rightarrow \infty$ to the initial function $f(x)$:

$$\lim_{n \rightarrow \infty} \max_{x \in [0,1]} |B_n[f](x) - f(x)| = \lim_{n \rightarrow \infty} \max_{x \in [0,1]} \Delta_n[f](x) = 0.$$

There exists a huge number of publications about the estimation of a value $\Delta_n[f](x)$, see for example [1], [4], [10], [11], [17], [18], [19], [23], [25], [30], [31], [32] and many others; see also reference therein.

Our purpose in this short report is obtaining a sharp up to multiplicative constant non-uniform and non-asymptotical universal estimate for the error of the Bernstein's approximate for the arbitrary initial continuous function in the terms of its module of continuity of the first order.

We intend to use too modern probabilistic apparatus: theory of Grand Lebesgue Spaces of random variables, in particular, theory of Strictly Subgaussian random variables; and introduce and use some generalization of this notion: Almost Strictly Subgaussian random variables.

Note that the non-uniform estimate are "much" better in the considered case than uniform ones, see [23], [24], [18] .

2 Auxiliary apparatus: subgaussian and strictly subgaussian random variables.

Let $\{\Omega, B, \mathbf{P}\}$ be some sufficiently rich probability space with expectation \mathbf{E} .

Definition 2.1.

We say that the *centered*: $\mathbf{E}\xi = 0$ numerical random variable (r.v.) ξ is subgaussian, or equally, belongs to the space $\text{Sub}(\Omega)$, if there exists some non-negative constant $\tau \geq 0$ such that

$$\forall \lambda \in R \Rightarrow \mathbf{E} \exp(\lambda \xi) \leq \exp[\lambda^2 \tau^2 / 2]. \quad (2.1).$$

The minimal value τ satisfying (2.1) for all the values $\lambda \in R$ is called a *subgaussian norm* of the variable ξ , write

$$\|\xi\|_{\text{Sub}} = \inf\{\tau, \tau > 0 : \forall \lambda \in R \Rightarrow \mathbf{E} \exp(\lambda \xi) \leq \exp(\lambda^2 \tau^2 / 2)\}.$$

Evidently,

$$\|\xi\|_{\text{Sub}} = \sup_{\lambda \neq 0} \left[\sqrt{2 \ln \mathbf{E} \exp(\lambda \xi) / |\lambda|} \right]. \quad (2.2)$$

This important notion was introduced by J.P.Kahane [14]; V.V.Buldygin and Yu.V.Kozachenko in [6] proved that the set $\text{Sub}(\Omega)$ relative the norm $\|\cdot\|$ is complete Banach space which is isomorphic to subspace consisting only from the centered variables of Orlicz's space over (Ω, B, P) with N – Orlicz-Young function $N(u) = \exp(u^2) - 1$; see also [16].

The detail investigation of this class of random variables with very interest applications into the theory of random fields reader may found in the book [8]; we reproduce here some main facts from this monograph.

Definition 2.1. The tail function $T_\eta(u)$, $u > 0$ for the numerical valued random variable η is as usually defined as follows

$$T_\eta(u) \stackrel{\text{def}}{=} \max[\mathbf{P}(\eta > u), \mathbf{P}(\xi < -u)], \quad u > 0.$$

If for instance $\|\xi\|_{\text{Sub}} = \tau \in (0, \infty)$, then

$$T_\xi(u) = \max[\mathbf{P}(\xi > u), \mathbf{P}(\xi < -u)] \leq \exp(-u^2/(2\tau^2)), \quad u \geq 0; \quad (2.3)$$

and the last inequality is in general case non-improvable. It is sufficient for this to consider the case when the r.v. ξ has the centered Gaussian non-degenerate distribution.

Conversely, if $\mathbf{E}\xi = 0$ and if for some positive finite constant K

$$T_\xi(u) = \max[\mathbf{P}(\xi > u), \mathbf{P}(\xi < -u)] \leq \exp(-u^2/K^2), \quad x \geq 0,$$

then $\xi \in \text{Sub}(\Omega)$ and $\|\xi\|_{\text{Sub}} < 4K$.

The subgaussian norm in the subspace of the centered r.v. is equivalent to the following Grand Lebesgue Space (GLS) norm:

$$\|\xi\|_{\text{GLS}} := \sup_{s \geq 1} \left[\frac{|\xi|_s}{\sqrt{s}} \right], \quad |\xi|_s = [\mathbf{E}|\xi|^s]^{1/s}.$$

More detail investigation of these spaces see in the monograph [20], chapter 1.

Denote in the sequel for brevity for any r.v. η

$$\sigma^2(\eta) = \sigma^2 = \text{Var } \eta = \mathbf{E}\eta^2 - (\mathbf{E}\eta)^2.$$

Definition 2.2. (See [8], chapter 1; [21].) The subgaussian r.v. ξ is said to be *Strictly Subgaussian*, write $\xi \in \text{SSub}$, iff

$$\forall \lambda \in R \Rightarrow \mathbf{E}e^{\lambda \xi} \leq e^{\lambda^2 \sigma^2(\xi)/2}, \quad (2.4)$$

or equally

$$\|\xi\|_{\text{Sub}} \leq \sigma(\xi) = \|\xi\|_{L_2(\Omega)}. \quad (2.4a)$$

Recall that always $\|\xi\|_{\text{Sub}} \geq \sigma(\xi) = \|\xi\|_{L_2(\Omega)}$, so that

$$\xi \in \text{SSub}(\Omega) \Leftrightarrow \mathbf{E}\xi = 0, \|\xi\| \text{Sub} = \sigma(\xi) = \|\xi\|_{L_2(\Omega)}. \quad (2.4b)$$

Many examples of strictly subgaussian distributions may be found in the book of V.V.Buldygin and Yu.V.Kozatchenko [8], chapter 1. For instance, arbitrary mean zero Gaussian distributed r.v. is strictly subgaussian, including the case when this r.v. is equal to zero a.e.; the symmetric Rademacher's r.v. ρ with distribution $\mathbf{P}(\rho = 1) = \mathbf{P}(\rho = -1) = 1/2$ belongs to the set $\text{SSub}(\Omega)$. The random variable η which has an uniform distribution on the symmetrical interval $(-b, b)$, $b = \text{const} \in (0, \infty)$ is Strictly Subgaussian.

Consider also for instance following the authors [8] the r.v. ζ with the following density:

$$f_\zeta(x) = \frac{\alpha + 1}{2\alpha} (1 - |x|^\alpha) I(|x| \leq 1), \quad \alpha = \text{const} \geq 0, \quad (2.5)$$

where $I(A) = I(A, x) = 1$, $x \in A$; $I(A) = I(A, x) = 0$, $x \notin A$ is indicator function; then $\zeta \in \text{SSub}(\Omega)$.

This example is interesting because the kurtosis of the r.v. ζ is zero if $\alpha = \sqrt{10} - 3$.

The convenience of these notions is following. Let $\{\xi(i)\}$, $i = 1, 2, \dots, n$ be (centered) independent subgaussian r.v. Denote

$$S(n) = \sum_{i=1}^n \xi(i), \quad \Sigma^2(n) = \sum_{i=1}^n (\|\xi(i)\| \text{Sub})^2. \quad (2.6)$$

Then $\|S(n)\| \text{Sub} \leq \Sigma(n)$ and following

$$\max(\mathbf{P}(S(n)/\Sigma(n) > x), \mathbf{P}(S(n)/\Sigma(n) < -x)) \leq e^{-x^2/2}, \quad x \geq 0, \quad (2.7)$$

the tail or concentrations inequalities.

If in addition $\xi(i)$ are identical distributed and $\beta := \|\xi(1)\| \text{Sub} \in (0, \infty)$, then

$$\sup_n \|S(n)/\sqrt{n}\| \text{Sub} = \beta$$

and

$$\sup_n \max(\mathbf{P}(S(n)/(\beta\sqrt{n}) > x), \mathbf{P}(S(n)/(\beta\sqrt{n}) < -x)) \leq e^{-x^2/2}, \quad x \geq 0, \quad (2.8)$$

If in addition the r.v. $\xi(i)$ are strictly subgaussian, the estimate (2.8) may be reinforced by lower estimate used the classical CLT:

$$\begin{aligned} \sup_n \mathbf{P}(S(n)/(\beta\sqrt{n}) > x) &\geq \lim_{n \rightarrow \infty} \mathbf{P}(S(n)/(\beta\sqrt{n}) > x) = \\ &= (2\pi)^{-1/2} \int_x^\infty e^{-y^2/2} dy \geq C x^{-1} e^{-x^2/2}, \quad x \geq 1. \end{aligned} \quad (2.9)$$

Definition 2.3. The non-degenerate centered random variable ν with variance σ^2 , $0 < \sigma < \infty$ is said to be *almost strictly subgaussian*, briefly $\nu \in \text{ASSub}$, if for all the real values $\lambda \in R$

$$\mathbf{E} \cosh(\lambda\nu/\sigma) \leq \exp(\lambda^2/2). \quad (2.10)$$

Proposition 2.1.

It follows immediately from the definition (2.10) the following tail estimate for these variables by means of Tchebychev-Chernov inequality

$$T_\nu(u) \leq 2 \exp(-u^2/2\sigma^2), \quad u \geq 0. \quad (2.11)$$

Proof. Obviously, if the r.v ν is almost strictly subgaussian: $\nu \in \text{ASSub}$ then $c \cdot \nu \in \text{ASSub}$. $c = \text{const}$.

Let now $\nu \in \text{ASSub}$ with $\sigma = 1$. We have by means of Tchebychev's inequality for the positive values λ and u :

$$\begin{aligned} T_\nu(u) &\leq \frac{e^{\lambda^2/2}}{\cosh(\lambda u)} = \frac{e^{\lambda^2/2}}{0.5(e^{\lambda u} + e^{-\lambda u})} \leq \\ &2 \exp(\lambda^2/2 - \lambda u) = 2 \exp(-u^2/2), \end{aligned}$$

if we choose $\lambda = u$.

Note that analogous approach appears at first in an article [22], where was applied to the investigation of the Central Limit Theorem in Banach space.

Denote also

$$\theta = \theta(p) = \sqrt{p \cdot (1 - p)}, \quad 0 < p < 1. \quad (2.12)$$

Let us now formulate and prove the main result of this section. Denote by $\mu = \mu_n = \mu_{n,p}$ the r.v. having Bernoulli (Binomial) distribution: $\text{Law } \mu = \text{Bin}(n, p)$, $n = 1, 2, \dots$; $0 < p < 1$. Recall that

$$\mathbf{P}(\mu = m) = \binom{n}{m} p^m (1 - p)^{n-m}, \quad m = 0, 1, 2, \dots, n$$

and

$$\mathbf{E}\mu = np, \quad \text{Var } \mu = n p (1 - p) = n \theta(p).$$

Remark 2.1. The non-centered random variable ζ may be named Almost Strictly Subgaussian iff it has a first moment and the centered r.v. $\zeta - \mathbf{E}\zeta$ satisfies the definition 2.3.

Theorem 2.1 (in our definitions and notations).

The random variables $\mu - np$, $0 < p < 1$, $n = 1, 2, \dots$ are almost strictly subgaussian uniformly for all the values n .

Proof. The inequality

$$\mathbf{E}(\mu - np)^{2m} \leq n^{-m} \frac{(2m)!}{2^m m!} \theta^m(p), \quad m = 0, 1, 2, \dots \quad (2.13)$$

is proved in the book of G.G.Lorentz [18], page 14. We get denoting the centered and normed r.v.

$$\eta := \frac{\mu - np}{\sqrt{np(1-p)}}$$

for all the real values $\lambda \in R$:

$$\begin{aligned} \mathbf{E} \cosh(\lambda \eta) &\leq \sum_{m=0}^{\infty} \frac{\lambda^{2m}}{(2m)!} \mathbf{E} \eta^{2m} \leq \\ \sum_{m=0}^{\infty} \frac{\lambda^{2m}}{(2m)!} \frac{(2m)!}{2^m m!} &= \sum_{m=0}^{\infty} \frac{\lambda^{2m}}{2^m m!} = e^{\lambda^2/2}. \end{aligned} \quad (2.14)$$

Remark 2.2. The inequality

$$\ln \mathbf{E} e^{\lambda \mu_1} \ln \left[(1-p)e^{-\lambda p} + pe^{\lambda(1-p)} \right] \leq p(1-p) \cdot \frac{\lambda^2}{2}, \quad (2.15)$$

where $1/2 \leq p \leq 1$ and $\lambda \geq 0$ is proved at first in an article of D.Berend and A.Kontorovich [2], lemma 5, page 4; see also some applications in [28] and a preliminary results in [7],[8], [15].

Perhaps, this form of the probabilistic inequality was known for the specialists in the approximation theory, see e.g. [13], [11], chapter 10, [23], [24].

3 Main result. Exactness.

Let us introduce the following (sublinear) operator (more precisely, the sequence of operators)

$$J_n[f](x) \stackrel{def}{=} \int_0^\infty \omega[f] \left(\frac{z \cdot \theta(x)}{\sqrt{n}} \right) z \exp \left(-\frac{z^2}{2} \right) dz.$$

Denote also

$$W = W(x) \stackrel{def}{=} \sup_{\text{const} \neq f \in C[0,1]} \sup_{n=1,2,\dots} \left[\frac{\Delta_n[f](x)}{J_n[f](x)} \right]. \quad (3.0)$$

Theorem 3.1. *We assert that for all the values $x \in (0, 1)$*

$$\frac{1}{\pi} \leq W = W(x) \leq 2. \quad (3.1)$$

Proof. Upper bound.

Let $x \in (0, 1)$; both the degenerate cases $x = 0$ or $x = 1$ are trivial.

The expression for $B_n[f](x)$ may be represented likewise in the initial work of S.N.Bernstein [3] as follows

$$B_n[f](x) = \mathbf{E}f(\mu/n), \quad (3.2)$$

where the r.v. μ has a binomial (Bernoulli) distribution with parameters $\mathbf{E}\mu = nx$, $\text{Var } \mu = nx(1-x)$. The r.v. $\mu - nx$ is equal to the sum

$$\mu - nx = \sum_{j=1}^n \xi_x(j), \Leftrightarrow \mu/n - x = \sum_{j=1}^n \xi_x(j)/n, \quad (3.3)$$

where the r.v. ξ_x has a binary distribution

$$\mathbf{P}(\xi_x = 1 - x) = x = 1 - \mathbf{P}(\xi_x = -x), \quad x \in (0, 1)$$

and the random variables $\xi_x(j)$ are independent copies of the r.v. ξ_x .

Put

$$\zeta = \zeta_{n,x} = \sum_{j=1}^n \xi_x(j)/\sqrt{n}, \quad (3.4)$$

then $\mu/n = x + \zeta_{n,x}/\sqrt{n}$. As we knew, the linear combination of strictly subgaussian independent random variables are also strictly subgaussian and we conclude therefore on the basis of theorem 2.1

$$T_{\zeta_{n,x}}(u) \leq 2 \exp\left(-u^2/2\theta^2(x)\right). \quad (3.5)$$

We can and will suppose without loss of generality that the r.v. $\zeta_{n,x}$ has a symmetrical distribution such that

$$T_{\zeta_{n,x}}(u) = 2 \exp\left(-u^2/2\theta^2(x)\right). \quad (3.5a)$$

More precisely, let $\tilde{\zeta} = \tilde{\zeta}_{n,x}$ be a symmetrically distributed r.v. such that

$$T_{\tilde{\zeta}_{n,x}}(u) = 2 \exp\left(-u^2/2\theta^2(x)\right), \quad (3.5b)$$

then for any positive monotonically increasing function $G = G(z)$, $z \geq 0$

$$\mathbf{E}G(\zeta_{n,x}) \leq \mathbf{E}G(\tilde{\zeta}_{n,x}).$$

We estimate consequently

$$\Delta_n[f](x) = |B_n[f](x) - f(x)| =$$

$$\left|\mathbf{E}(f(x + \zeta/\sqrt{n}) - f(x))\right| \leq \mathbf{E}\omega[f]\left(\frac{\zeta_{n,x}}{\sqrt{n}}\right),$$

and we deduce after integration by parts relying on the inequality (3.5)

$$\Delta_n[f](x) \leq 2 \int_0^\infty \omega[f]\left(\frac{y}{\sqrt{n}}\right) \frac{y}{\theta^2(x)} \exp\left(-y^2/2\theta^2(x)\right) dy =$$

$$2 \int_0^\infty \omega[f] \left(\frac{z\theta(x)}{\sqrt{n}} \right) z \exp(-z^2/2) dz = 2 J_n[f](x). \quad (3.6)$$

Thus, we proved the upper bound: $W \leq 2$.

Remark 3.1 Note that $\theta(x) \leq 1/2$, therefore

$$\Delta_n[f](x) \leq 2 \int_0^\infty \omega[f] \left(\frac{y}{2\sqrt{n}} \right) y \exp(-y^2) dy. \quad (3.6a)$$

Thus, we obtained easily the uniform Bernstein's error estimate through the uniform one.

Remark 3.2 It follows immediately from the estimation (3.6a) and on the basis of Lebesgue dominated convergence theorem that

$$\lim_{n \rightarrow \infty} \max_{x \in [0,1]} \Delta_n[f](x) = 0,$$

i.e. the classical Bernstein's result.

Examples 3.1.

Let the function $f = f(x)$ be Hölderian:

$$\exists \alpha = \text{const} \in (0, 1], \quad H_\alpha = \text{const} < \infty \Rightarrow$$

$$|f(x_1) - f(x_2)| \leq H_\alpha |x_1 - x_2|^\alpha, \quad x_1, x_2 \in [0, 1]; \quad (3.7)$$

or equally

$$\omega[f](\delta) \leq H_\alpha \delta^\alpha, \quad H_\alpha < \infty. \quad (3.7a)$$

It follows immediately from theorem 3.1

$$\Delta_n[f](x) \leq 2 H_\alpha \left[\frac{2x(1-x)}{n} \right]^{\alpha/2} \Gamma\left(\frac{\alpha}{2}\right). \quad (3.8)$$

If in addition $\alpha = 1$ (Lipshitz condition) and if we denote as ordinary $L = H_1 \in (0, \infty)$, then

$$\Delta_n[f](x) \leq 2 L \left[\frac{2\pi x(1-x)}{n} \right]^{1/2}. \quad (3.8a)$$

See also [19], [25].

Lower bound.

Ranko Bojanic in the article [4], see also [5], [31], [30] considered the family of a trial functions of a form

$$g_t(x) = |t - x|, \quad t, x \in [0, 1],$$

which is Lipschitzian relative to the variable x , (as well as relative to the variable t), with constant $L = 1$, and proved in fact that as $n \rightarrow \infty$

$$\Delta_n[g_t(\cdot)] = \left[\frac{2x(1-x)}{\pi n} \right]^{1/2} + o\left(\frac{1}{n}\right). \quad (3.9)$$

We conclude therefore that for all the values x from the (open) interval $x \in (0, 1)$

$$W(x) \geq \lim_{n \rightarrow \infty} \left[\frac{\Delta_n[g_t](x)}{J_n[g_t](x)} \right] = 1/\pi. \quad (3.10)$$

This completes the proof of theorem 3.1.

4 Convergence of derivative for Bernstein's polynomials.

Let now in this section the aforementioned function f be continuous differentiable: $f \in C^1[0, 1]$. It is known that the derivatives $B'_n[f](x)$ converges uniformly to one for the function $f'(x)$, see [12], [33].

We intend to obtain here the refined estimation also non-asymptotical and non-uniform for the approximation

$$\Delta_n^1[f](x) \stackrel{def}{=} \max_x |B'_n[f](x) - f'(x)|. \quad (4.1)$$

Theorem 4.1.

$$\begin{aligned} \Delta_n^1[f](x) &\leq \frac{3}{2} \omega[f']\left(\frac{1}{n}\right) + 2 J_{n-1}[f'](x) = \\ &\frac{3}{2} \omega[f']\left(\frac{1}{n}\right) + 2 \int_0^\infty \omega[f']\left(\frac{z \cdot \theta(x)}{\sqrt{n-1}}\right) z \exp\left(-\frac{z^2}{2}\right) dz, \quad n \geq 2. \end{aligned} \quad (4.2)$$

Proof.

Let $n \geq 2$. The expression for $B'_n[f](x)$ is given, for example, in the article [33]:

$$B'_n[f](x) = \sum_{j=0}^{n-1} n V_{1/n}[f]\left(\frac{j}{n}\right) \binom{n-1}{j} x^j (1-x)^{n-1-j}, \quad (4.3)$$

where

$$V_{1/n}[f]\left(\frac{j}{n}\right) = f\left(\frac{j+1}{n}\right) - f\left(\frac{j}{n}\right)$$

is a difference operator for the function f with step $1/n$ at the point j/n .

Note that

$$\left| n V_{1/n}[f] - f' \left(\frac{j}{n} \right) \right| \leq \frac{1}{2} \omega[f'] \left(\frac{1}{n} \right). \quad (4.4)$$

Further, since $n \geq 2$

$$\left| f' \left(\frac{j}{n} \right) - f' \left(\frac{j}{n-1} \right) \right| \leq \omega[f'] \left(\frac{1}{n} \right),$$

therefore (see theorem 3.1)

$$\begin{aligned} \Delta_n^1[f](x) &\leq \frac{3}{2} \omega[f'] \left(\frac{1}{n} \right) + | B_{n-1}[f'](x) - f'(x) | \leq \\ &\frac{3}{2} \omega[f'] \left(\frac{1}{n} \right) + 2 J_{n-1}[f'](x), \end{aligned} \quad (4.5)$$

Q.E.D.

Remark 4.1.

As for the lower estimate for Bernstein derivative estimate, consider the trial functions of the form

$$G_t = G_t(x) := \int_0^x g_t(y) dy; \quad (4.6)$$

then for all the values $t \in [0, 1]$

$$\Delta_n^1[G_t](x) \geq \int_0^\infty \omega[G'_t(x)] \left(\frac{z \cdot \theta(x)}{\sqrt{n}} \right) z \exp(-z^2/2) dz. \quad (4.7)$$

5 Multivariate Bernstein's polynomials.

The speed of convergence.

Let now $f = f(x, y)$, $(x, y) \in [0, 1]^2$ be continuous numerical function defined on the square $[0, 1]^2$, $n_1, n_2 \geq 2$ be integer numbers.

One can define the modulus of continuity of the multivariate function $f = f(x, y)$ as follows

$$\omega[f](\delta_1, \delta_2) \stackrel{def}{=} \sup_{|x_1 - y_1| \leq \delta_1} \sup_{|x_2 - y_2| \leq \delta_2} |f(x_1, y_1) - f(x_2, y_2)|, \quad \delta_{1,2} \in [0, 1]. \quad (5.0)$$

The multivariate (more exactly, bivariate) Bernstein's polynomial $B_{n_1, n_2}[f](x, y)$ of an order (n_1, n_2) one can define by a formula

$$B_{n_1, n_2}[f](x, y) \stackrel{def}{=}$$

$$\sum_{k_1=0}^{n_1} \sum_{k_2=0}^{n_2} \binom{n_1}{k_1} \binom{n_2}{k_2} f \left(\frac{k_1}{n_1}, \frac{k_2}{n_2} \right) x^{k_1} (1-x)^{n_1-k_1} y^{k_2} (1-y)^{n_2-k_2}, \quad (5.1)$$

see, e.g. [18], [33].

We define formally by the analogy with the one-dimensional case

$$B_{n_1, \infty}[f](x, y) \stackrel{\text{def}}{=} \sum_{m=0}^{n_1} \binom{n_1}{m} f\left(\frac{m}{n_1}, y\right) x^m (1-x)^{n_1-m} \quad (5.1a)$$

and analogously

$$B_{\infty, n_2}[f](x, y) \stackrel{\text{def}}{=} \sum_{k=0}^{n_2} \binom{n_2}{k} f\left(x, \frac{k}{n_2}\right) y^k (1-y)^{n_2-k} \quad (5.1b).$$

Finally,

$$B_{\infty, \infty}[f](x, y) \stackrel{\text{def}}{=} f(x, y), \quad f(\cdot, \cdot) \in C([0, 1]^2).$$

Define also the two-dimensional non-uniform error of Bernstein approximation

$$\Delta_{n_1, n_2}[f](x, y) := | B_{n_1, n_2}[f](x, y) - f(x, y) |. \quad (5.2)$$

Introduce analogously to the one-dimensional case the following operator (more precisely, the sequence of operators) $J_{n_1, n_2}[f](x, y) \stackrel{\text{def}}{=}$

$$\int_0^\infty \int_0^\infty \omega[f]\left(\frac{z_1 \cdot \theta(x)}{\sqrt{n_1}}, \frac{z_2 \cdot \theta(y)}{\sqrt{n_2}}\right) z_1 z_2 \exp\left(-\frac{z_1^2}{2} - \frac{z_2^2}{2}\right) dz_1 dz_2. \quad (5.3)$$

Theorem 5.1. *We propose as before that for all the values $x, y \in [0, 1]$ and $d = 2$*

$$\pi^{-1} \leq W_2 = W_2(x, y) \stackrel{\text{def}}{=} \sup_{\text{const} \neq f \in C[0, 1]^2} \sup_{n_1, 2=1, 2, \dots} \left[\frac{\Delta_{n_1, n_2}[f](x, y)}{J_{n_1, n_2}[f](x, y)} \right] \leq 2^d. \quad (5.4)$$

Proof.

Our reasoning are likewise ones in the section 3. Namely, let $0 < x, y < 1$; we have

$$B_{n_1, n_2}[f](x, y) = \mathbf{E} f\left(\frac{\mu_1}{n_1}, \frac{\mu_2}{n_2}\right), \quad (5.5)$$

where the *independent* random variables μ_1 and μ_2 have the integer binomial (Bernoulli) distribution with the parameters correspondingly

$$\mathbf{E}\mu_1 = n_1 x, \quad \mathbf{E}\mu_2 = n_2 y, \quad \text{Var } \mu_1 = n_1 x(1-x), \quad \text{Var } \mu_2 = n_2 y(1-y). \quad (5.6)$$

Each such the variable may be represented as above

$$\frac{\mu_1}{n_1} = x + \frac{\zeta_{1,x}}{\sqrt{n_1}}, \quad \frac{\mu_2}{n_2} = y + \frac{\zeta_{2,y}}{\sqrt{n_2}}, \quad (5.7)$$

where both the r.v. $\zeta_{1,x}$ and $\zeta_{2,y}$ are independent and are as above almost strictly subgaussian.

We can and will suppose without loss of generality that the r.v. $\zeta_{1,x}$, $\zeta_{2,y}$ have a symmetrical distribution such that

$$T_{\zeta_{1,x}}(u_1) = \exp\left(-u_1^2/2\theta^2(x)\right), \quad (5.8a)$$

$$T_{\zeta_{2,y}}(u_2) = \exp\left(-u_2^2/2\theta^2(y)\right) \quad (5.8b)$$

and are independent.

It is no hard to calculate as above the densities of both the independent variables $\zeta_{1,x}$, $\zeta_{2,y}$.

We estimate consequently

$$\Delta_{n_1,n_2}[f](x,y) = |B_{n_1,n_2}[f](x,y) - f(x,y)| = 2^d \times$$

$$|\mathbf{E}(f(x + \zeta_{1,x}/\sqrt{n_1}, y + \zeta_{2,y}/\sqrt{n_2}) - f(x,y))| \leq 2^d \mathbf{E}\omega[f]\left(\frac{\zeta_{1,x}}{\sqrt{n_1}}, \frac{\zeta_{2,y}}{\sqrt{n_2}}\right),$$

and we deduce after integration twice by parts analogously to the proof of the inequality (3.5) $\Delta_{n_1,n_2}[f](x,y) \leq 2^d \times$

$$\begin{aligned} & \int_0^\infty \int_0^\infty \omega[f]\left(\frac{v_1}{\sqrt{n_1}}, \frac{v_2}{\sqrt{n_2}}\right) \frac{v_1}{\theta^2(x)} \frac{v_2}{\theta^2(y)} \exp\left(-v_1^2/2\theta^2(x) - v_2^2/2\theta^2(y)\right) dv_1 dv_2 = \\ & 2^d \cdot \int_0^\infty \int_0^\infty \omega[f]\left(\frac{z_1 \cdot \theta(x)}{\sqrt{n_1}}, \frac{z_2 \cdot \theta(y)}{\sqrt{n_2}}\right) z_1 z_2 \exp\left(-\frac{z_1^2}{2} - \frac{z_2^2}{2}\right) dz_1 dz_2 = \\ & 2^d J_{n_1,n_2}[f](x,y). \end{aligned} \quad (5.9)$$

Thus, we proved the upper bound: $W_2 \leq 2^d$.

To derive the lower bound, it is sufficient to consider as the capacity of an example the *factorable* function, namely

$$f_0(x,y) = g_{t_1}(x) \cdot h(y), \quad t_1 \in (0,1), \quad h(\cdot) \in C[0,1],$$

so that

$$B_{n_1,n_2}[f_0](x,y) = B_{n_1}[g_{t_1}](x) \cdot B_{n_2}[h](y);$$

$$\omega[f_0](\delta_1, \delta_2) = \omega[g_{t_1}](\delta_1) \cdot \omega[h](\delta_2),$$

and finally

$$J_{n_1,n_2}[f_0](x,y) = J_{n_1}[g_{t_1}](x) \cdot J_{n_2}[h](y).$$

Let us choose, for example, $h(y) = \text{const} = 1$. It follows immediately from the second proposition of theorem 3.1

$$\begin{aligned} W_2(x, y) &\geq \lim_{n_1 \rightarrow \infty} \frac{\Delta_{n_1, \infty}[f_0](x, y)}{J_{n_1, \infty}[f_0](x, y)} = \\ \lim_{n_1 \rightarrow \infty} \frac{\Delta_{n_1}[g_{t_1}](x)}{J_{n_1}[g_{t_1}](x)} &= W(x) = 1/\pi, \end{aligned} \quad (5.10)$$

Q.E.D.

More general approach.

Let $||(x, y)||$ be any non-degenerate norm on the plane R^2 , for instance,

$$||(x, y)|| = \sqrt{x^2 + y^2}, \quad ||(x, y)|| = \max(|x|, |y|), \quad ||(x, y)|| = |x| + |y| \quad (5.11)$$

and so one.

In accordance with this definition suppose that there exists non-negative continuous non-decreasing function $\gamma(\delta_1, \delta_2) = \gamma[f](\delta_1, \delta_2)$ such that $\gamma(0+, 0+) = 0$ and

$$|f(x_1, y_1) - f(x_2, y_2)| \leq \gamma[f](||(x_1 - x_2, y_1 - y_2)||).$$

We deduce as above $\Delta_{n_1, n_2}[f](x, y) \leq 2^d \times$

$$\int_0^\infty \int_0^\infty \gamma[f] \left(\frac{v_1}{\sqrt{n_1}}, \frac{v_2}{\sqrt{n_2}} \right) \frac{v_1}{\theta^2(x)} \frac{v_2}{\theta^2(y)} \exp \left(-v_1^2/2\theta^2(x) - v_2^2/2\theta^2(y) \right) dv_1 dv_2. \quad (5.12)$$

6 Concluding remarks.

It is no hard to generalize the results of the last section into the "more" multivariate case $d = 3, 4, 5, \dots$; as well as into other methods of approximation, if only they had a probabilistic representation.

One can also investigate and improve the rate of convergence of partial derivatives for the multivariate Bernstein's polynomials.

References

- [1] ACHIEZER (AKHIEZER) N.I. *Theory of approximation*. Frederick Ungar Publishing Co., New York, 1956.

- [2] BEREND D. AND KONTOROVICH A. *On the concentration of the missing mass.* Electron. Commun. Probab., 18, (3): 1-7, 2013;
arXiv:1210.3248v1 [math.PR] 11 Oct 2012.
- [3] BERNSTEIN S.N. *Demonstration du theoreme de Weierstrass, fonde sur le probabilites.* Comm. Sot. Math. Kharkow, **13**, (1912-1913), 1-2.
- [4] BOJANIC RANKO. *On the approximation of continuous functions by Bernstein polynomials.* Acad. Serbe Sci. Arts Glus, 232, (1959), 59-65. [In Serbian].
- [5] BOJANIC RANKO AND FUHUA CHENG. *Rate of Convergence of Bernstein Polynomials for Functions with Derivatives of Bounded Variation.* Journal of Mathematical Analysis and Applications. **141**, 136-151, (1989).
- [6] BULDYGIN V.V., KOZATCHENKO YU.V. *About subgaussian random variables.* Ukrainian Math. Journal, 1980, **32**, N^o 6, 723-730.
- [7] BULDYGIN V.V., MOSKVICHOVA K.K. *The sub - Gaussian norm of a binary random variable.* Theor. Probability and Math. Statist. Vip. 86, No. 86, 2013, Pages 33-49.
- [8] BULDYGIN V.V., KOZATCHENKO YU.V. *Metric Characterization of Random Variables and Random Processes.* 1998, Translations of Mathematics Monograph, AMS, v.188.
- [9] S. X. CHEN AND J. S. LIU. *Statistical applications of the Poisson-binomial and conditional Bernoulli distributions.* Statist. Sinica, 7(4): 875-892, 1997.
- [10] CHLODOVSKY, I. N. *On some properties of S. N. Bernstein polynomials.* Proc. 1st All-Union congress of mathematics (Kharkov, 1930), Moscow; Leningrad: ONTI NKTP USSR, 1936, 22, (in Russian).
- [11] DeVORE, LORENTZ G.G. *Constructive approximation.* Springer Verlag, 1997.
- [12] FLOATER M. S. *On the convergence of derivatives of Bernstein approximation.* Journal of Approximation Theory, 2005, 134(1), 130135.
- [13] GONSKA H., ZHOU D-X. *On an extremal problem concerning Bernstein operators,* Serdica Math. J. 21, 1995, 137-150.
- [14] KAHANE J.P. *Proprieties locales des fonctions a series de Fourier aleatoires.* Studia Math. (1960), **19**, N^o 1, 1-25.
- [15] KEARNS M. AND SAUL L. *Large deviation methods for approximate probabilistic inference.* In Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, pages 311-319. Morgan Kaufmann Publishers Inc., 1998.
- [16] KOZATCHENKO YU.V., OSTROVSKY E.I. *Banach spaces of random variables of subgaussian type.* Theory Probab. And Math. Stat., Kiev, (1985), p. 42-56 (in Russian).

- [17] LI Z. *Bernstein polynomials and modulus of continuity*. J. Approx. Theory, **102**, (2000), 171-174; MR1736050 (2000j:41029).
- [18] LORENTZ G.G. *Bernstein Polynomials*. Univ. of Toronto Press, Toronto, 1953.
- [19] MATH'É P. *Approximation of Hölder continuous functions by Bernstein polynomials*. Amer. Math. Monthly, **106**, (1999), 568-574.
- [20] OSTROVSKY E.I. *Exponential Estimations for Random Fields*. Monograph, Moscow-Obninsk, OINPE, (1999), (in Russian).
- [21] OSTROVSKY E., SIROTA L. *Subgaussian and strictly subgaussian random variable*. arXiv:1406.3933v1 [math.PR] 16 Jun 2014
- [22] OSTROVSKY E.I. *Generalization of the norm Buldygin - Kosatchenko and Central Limit Theorem in Banach space*. Probab. Theory Appl., (1982), issue 3, p. 616-618. (in Russian).
- [23] PĂLATĂNEA RADU. *On some constants in approximation by Bernstein operators*. General Mathematics, Vol. 16, No. 4, (2008), 137-148.
- [24] PĂLATĂNEA RADU. *Optimal constant in approximation by Bernstein operators*. J. Comput. Analysis Appl., 5 (2), 2003, 195-235.
- [25] PRESTIN J. *Best approximation in Lipschitz space*. Coll. Math. Soc. J. Bolyai, **49**, (1985), 753-759.
- [26] RAGINSKY M. AND SASON I. *Concentration of Measure Inequalities in Information Theory, Communications, and Coding*. Foundations and Trends in Communications and Information Theory, vol. 10, no. 1-2, pp. 1246, 2013.
- [27] RIVASPLATA OMAR. *Subgaussian random variables: An expository note*. Internet publication, PDF, November 12, 2012.
- [28] SCHLEMM E. *The Kearns-Saul inequality for Bernoulli and Poisson-binomial distributions*. arXiv:1405.4496v1 [math.PR] 18 May 2014
- [29] SEROV A.A., ZUBKOV A.M. *A full proof of universal inequalities for the distribution function of the binomial law*. arXiv:1207.3838v1 [math.PR] 16 Jul 2012
- [30] TELYAKOVSKII S. A. *On the rate of approximation of functions by the Bernstein polynomials*. Proc. Inst. Math. Mech., 2009, 264, suppl. 1, 177184.
- [31] TIKHONOV, I. V., SHERSTYUKOV, V. B. *The module function approximation by Bernstein polynomials*. Cheliabinsk State Univ. Vestnik, Ser. Matem., Mech., Inform., 2012, **15**, (26), 6-40. (in Russian).
- [32] TIMAN A.F. *Theory of Approximation of Functions of a Real Variable*. International Series of Monographs on Pure and Applied Mathematics, Paperback-January 1, 1963. Pergamon Press.

- [33] VERETENNIKOV A.YU., VERETENNIKOVA E.V. *On convergence of partial derivatives of multivariate Bernstein polynomials.*
arXiv:1507.05235v1 [math.PR] 18 Jul 2015
- [34] ZUBKOV A.M., SEROV A.A. *Bounds for the number of Boolean functions admitting affine approximations of a given accuracy.* Discrete Math. Appl., 2010, 20, N^o 5-6, p. 467-486.